# Appendix to The Tularosa Study: An Experimental Design and Implementation to Quantify the Effectiveness of Cyber Deception

Kimberly Ferguson-Walter
U.S. Department of Defense
Washington, D.C.

Temmie Shade
U.S. Department of Defense
Washington, D.C.

Andrew Rogers
U.S. Department of Defense
Washington, D.C.

Mike Trumbo
Sandia National Laboratories
Albuquerque, New Mexico

Kevin Nauer
Sandia National Laboratories
Albuquerque, New Mexico

Kristin Divis
Sandia National Laboratories
Albuquerque, New Mexico

Aaron Jones
Sandia National Laboratories
Albuquerque, New Mexico

Angela Combs
Sandia National Laboratories
Albuquerque, New Mexico

Robert Abbott
Sandia National Laboratories
Albuquerque, New Mexico

## A INDIVIDUAL MEASURES

### A.1 Red Team Briefing

Following the network penetration task each day, participants were asked to spend 15 minutes responding to an open-ended question about their experience. The following language was used to prompt participants, with the day updated to "ONE" or "TWO" and the underlined portion (not underlined for participants) designating this was only displayed to participants who were in an informed condition that day:

Please take 15 minutes to brief us on your experience during the cyber task on DAY ONE (today). Please share any information you think is relevant or important for a briefing. Specific questions to consider include: major vulnerabilities found, flaws in the network, success in exfiltrating assets, strategies you used, aspects of the network that were particularly frustrating and/or confusing, and nature of deception on network, if found.

### A.2 Overall Briefing

Following the Day 2 Red Team Briefing, participants were to respond the following open-ended question about their experience:

Please take 15 minutes to brief us on your overall experience during the cyber tasks across BOTH DAYS (today and yesterday). Please share any information you think is relevant or important for a briefing. Specific questions to consider include: information not included in either daily briefing, changes in strategy or approach between the days, differences noted between the days, suspicions about the networks, etc.

They were then asked to answer the following questions:

How much do you rely on each source of information/reference material during a typical engagement on a scale from 1 to 5? (With "1" indicating not at all and "5" indicating frequently).

- Public Internet (website/forums)
- Corporate forums (e.g., internal wiki)
- Professional network (friends/colleagues)
- Private forums (e.g., restricted IRC channel)
- Personal resources (e.g., code repositories, notes)
- Books/printed materials

How would you rate the tools available to you a scale from 1 to 5? (With "1" indicating none of the tools you needed were available and "5" indicating you had every tool you needed).

Were there any tools you would normally rely on that we didn?t give you? If so, which ones?

Before coming to participate in this exercise, did you do any research on the project beyond the information provided in the recruitment message? If so, please describe.

Did you discuss the cyber task with other red teamers (e.g., at lunch or between Day 1 and Day 2)? If so, what did you talk about?

### A.3 Cyber Task Questionnaire

On each day, participants were asked about the psychological and cognitive effects of their experience during the network penetration task. The following language was used to prompt participants, with the day updated to "ONE" or "TWO" and the underlined portion (not underlined for participants) only displayed to participants on Day 2:

While working on the cyber task on DAY ONE:

(1) On a scale from 1-5, how much **confusion** did you experience throughout the task? (With "1" indicating

you were never confused and "5" indicating you were always confused). What caused your confusion?
(2) On a scale from 1-5, how much **self-doubt** did you experience throughout the task? (With "1" indicating you never doubted yourself and "5" indicating you were always doubting yourself). What caused your self-doubt?
(3) On a scale from 1-5, how **confident** did you feel throughout your attack? (With "1" indicating not confident at all and "5" indicating very confident).
(4) On a scale from 1-5, how **surprised** were you during the task by unexpected aspects of the network? (With "1" indicating not at all surprised and "5" indicating very surprised). What surprised you?
(5) On a scale from 1-5, how **frustrated** were you during the task by unexpected aspects of the network? (With "1" indicating not at all frustrated and "5" indicating very frustrated). What frustrated you?
(6) Please describe your planned, attempted, successfully executed, and/or unsuccessfully executed **strategies**.
(7) Do you believe **deception** was present on the network on either Day 1 or Day 2? If so, what do you believe the deception entailed? On which day or days was it present?

## A.4 Demographics Questionnaire

Participants were asked to answer the following questions:

What is your gender?

- Male
- Female
- Other

What is your age range?

- Less than 35 years
- 35-50 years
- Over 50 years

What is the highest level of education you've completed?

- High School
- Associates/Technical School
- Bachelors
- Masters
- PhD

Is English your primary language?

- English is primary language
- English is secondary language

## A.5 Experience Questionnaire

Participants were asked the following questions about their red teaming experience:

For each of the following areas, please rate your level of expertise on a scale of 1 to 5 (1 = novice, 5 = expert):

- Cyber security
- Network penetration
- Host penetration
- Network reconnaissance
- Incidence response
- Generalized defense practice
- Network protocol reverse engineering
- Binary reverse engineering

How involved are you in each phase of an engagement, on a scale of 1 to 5 (1 = least, 5 = most)? (Phases from Lockheed Martin "Cyber Kill Chain").

- Reconnaissance (e.g., harvesting email addresses)
- Weaponization (coupling exploit with backdoor into deliverable payload)
- Delivery of weaponized bundle via email, web, USB, etc.
- Exploitation (execute code on victim?s system)
- Installation of malware on the asset
- Command and control channel for remote manipulation of the victim
- Actions on objectives/accomplishment of goals

How well do each of these objectives describe a typical engagement you are involved with, on a scale of 1 to 5 (1 = least, 5 = most)?

- Compliance testing (e.g., HPPA)
- Blue team training
- Demonstrate needs for increased security investments
- Whiteboarding / gaming / tabletop exercises
- Post-attack remediation effort
- Vulnerability analysis (e.g., source code / reverse engineering)
- Security architecture review
- Persistent adversary (APT) emulation

Please indicate how many years of experience you have in each of the following areas:

- Cyber security
- Network penetration
- Host penetration
- Network reconnaissance
- Incidence response
- Generalized defense practice
- Network protocol reverse engineering
- Binary reverse engineering

Which operating system do you use the most (Linux, Windows, or Other)? If "Other" please specify.

What is the context in which you generally work? Please answer each of the following:

- Size of the team you normally work in (Individually, 2-3 people, or 4 or more people)
- What is the total duration of a typical engagement (1-2 days, 3 days-1 week, 1-2 weeks, 2 weeks to one month, or over one month)?
- Types of expertise on the team (place an X next to each category, as applies to the core team):
  – Network penetration

  - Host penetration
  - Network reconnaissance
  - Incidence response
  - Generalized defense practice
  - Network protocol reverse engineering
  - Binary reverse engineering
  - Other (Please Specify)
- Expertise of other people you have easy access to, if needed (place an X next to each that applies):
  - Network penetration
  - Host penetration
  - Network reconnaissance
  - Incidence response
  - Generalized defense practice
  - Network protocol reverse engineering
  - Binary reverse engineering
  - Other (Please Specify)

### A.6 Deception Questionnaire

Participants were asked the following open-ended questions:

- What makes you suspicious?
- When you experience something as suspicious, what do you interpret it as?
- When attacking a system, would you be likely to you think that the system has deception mechanisms in place?
- When attacking a system, do you first look for signs for deception?
- How do you respond when you suspect deception is in the system?
- How do you respond when you confirm the system is utilizing deception?
- If you attacked a system where deception was used, how likely are you to think deception will be present the next time you attack it?
- If you attacked a system where deception was used, how likely is it that you will attack the system again?
- If you attacked a system where deception was used, do you think that a Blue Team is also operating as part of the defense?
- If the system explicitly warned you that deception is present, how likely are you to believe the message?
- If we wanted to convince attackers that deception is present, what should we do?
- If we wanted to convince attackers that no deception is present, what should we do?

## B TASK BRIEFING

See below for the exact wording used in the task briefing at the start of the day. The underlined sections were only shown to participants in the informed condition (and were not underlined for participants).

**Scenario**

You represent an APT group attempting to gather information from the company Demokratika Petroleum (abbreviated as DP). You have achieved an initial foothold on the DP company network, and now must discover as much as you can about potentially valuable targets on the network. You will conduct recon on the network and locate vulnerable services, misconfigurations, and working exploits. Specifically, your task is to provide actionable intelligence about the company network which can be used by the follow-on team over the next 3-6 months. Your objective is to collect as much relevant information about the target network as you can in the allotted time without compromising future network operations.

There may be deception on the network.

**Procedures**

(1) You will access the DP network using a dedicated laptop which has a Kali Linux operating system to use for reconnaissance and system exploitation (user: root password: toor). There is a Kali repository installed on the computer and you may install additional tools as needed during your activities.

(2) You will also have access to a second laptop which is connected to the internet for research and technical assistance (user: recoilforce password: f0r3ns1c). However, you may not electronically transfer information from this internet connected laptop to the attack laptop (or vice versa); you must manually enter all commands, reporting, etc.

(3) When you learn potentially useful information about target systems on this network you will immediately report this information to your team via your internet connected laptop using the Mattermost website at mattermost-dev.recoilforce.net using the following format:
   - The last 2 octets of the IP address
   - Why you believe the host is interesting
   - How you obtained this information
   - Estimate its value to future operations

   You don't need to be sure about a host to file a report; you can make multiple reports on the same host. Normally you will not receive a reply to these reports, but **they are your primary deliverable**.

(4) Additional notes, commands, etc (that are not sent in a Mattermost report) should be kept in the file /root/notes

(5) We will be monitoring your progress, and taking into account how noisy your activities are. Prioritize obtaining as much actionable intelligence about target systems as possible without compromising future operations on the target network.

(6) If you experience any technical difficulties, you can reach technical support using Mattermost at mattermost-dev.recoilforce.net, which is the homepage in Firefox.

(7) A proctor will be present for general questions, including help contacting technical support. The proctors and tech support are not role-players in the simulation and may not be consulted for help in performing

tasks on the network; they are here to facilitate your independent effort.

(8) If you need to reboot either laptop for any reason, ask a proctor for assistance so that we can ensure it is collecting the data for this exercise. (For example, the attack laptop is running screen capture and keyboard capture programs).

**Ground Rules:**

- Limit your recon/attacks to the simulation network, 192.168.5.0/24. Within this network, do not perform attacks against the NTP server, located at 192.168.5.2 (it provides accurate time for data collection purposes and is not relevant to the task). The DP infrastructure is virtualized. You may not attack the virtual infrastructure (the hypervisor). You may not perform physical attacks on the system or social engineering attacks.
- Do not stop the recording programs running on our laptops (e.g. screen and keyboard capture). The information collected is important to the exercise we have hired you to support and will not be linked to your identity. Please help us protect your privacy by NOT entering any personally-identifying information (such as using your name in your notes or Mattermost reports, or logging into Facebook) on either laptop.
- You may not make copies of information (including software) from any of our computer systems to any storage device or computer system except the ones we have provided. Do not enable the WiFi on the attack client computer or connect it to any network other than the simulation network provided.
- Do not disclose your observations about the network simulation, its vulnerabilities, or defenses encountered. This includes not discussing your observations with other participants present at this event or with individuals that might be participating in future sessions; each individual's performance must be independent. This is important to the scientific validity of our results.
- You are expected to utilize your cyber-security subject matter expertise and perform to the best of your ability, however you are not required to utilize knowledge or techniques deemed proprietary by your employer.

# C SCHEDULE

**Day 1**

*8:30 A.M. to 9:00 A.M. (Introduction and Set-Up)*: Participants were introduced to the study and assigned a work station. Participants who opted in to the HSR portion also had the Empatica E4 set up and filled out the Experience Questionnaire. All participants worked through an electronic task briefing to orient themselves with the red teaming scenario (see Appendix B). Those in the informed condition were also verbally informed that deception may be present on the network.

*9:00 A.M. to 11:30 A.M. (Cyber Task, Part 1)*: Participants started on the network penetration task. Proctors noted the timing of breaks and any extreme behaviors (e.g., slamming mouse down in frustration) in the HSR subjects.

*11:30 A.M. to 12:00 P.M. (Lunch Break)*: Participants were given a lunch break and reminded not to discuss the details of the cyber task, as per the nondisclosure agreement.

*12:00 P.M. to 4:00 P.M. (Cyber Task, Part 2)*: Participants continued the network penetration task. Proctors continued to note the timing of breaks and any extreme behaviors in HSR subjects.

*4:00 P.M. to 4:15 P.M. (Briefing)*: All activity on the attack laptops was halted and participants filled out the Day 1 Red Team Briefing (see Appendix A)

*4:15 P.M. to 5:15 P.M. (Task Battery or Report Writing)*: Participants who opted out of the HSR portion continued to write a report on the cyber task (continuing the red team briefing). Participants who opted into the HSR portion complete the following tasks in order: Shipley-2 (hard copy), Day 1 Cyber Task Questionnaire (hard copy), Demographics Questionnaire (computer), Big Five Inventory (computer), General Decision-Making Style Inventory (computer), Indecisiveness Scale (computer), Sandia Matrices (computer), Over-Claiming Questionnaire (computer), and Sleep Quality Questionnaire (computer).

*5:15 P.M. to 5:30 P.M. (Wrap-Up)*: Participants were reminded what to expect the next day and not to discuss the task with others. Proctors collected the Empatica E4 devices from participants who participated in the HSR portion of the study.

**Day 2**

*8:30 A.M. to 9:00 A.M. (Introduction and Set-Up):* Participants were reminded of the rules of engagement and told they would be working on a separate network on Day 2 (compared to Day 1). Participants who opted into the HSR portion of the study also had the Empatica E4 devices set up. All participants were given a hard copy of the task briefing document (see Appendix B); those in the informed condition were verbally told that deception may be present on the network.

*9:00 A.M. to 11:30 A.M. (Cyber Task, Part 1):* Participants started on the network penetration task. Proctors noted the timing of breaks and any extreme behaviors (e.g., slamming mouse down in frustration) in the HSR subjects.

*11:30 A.M. to 12:00 P.M. (Lunch Break):* Participants were given a lunch break and reminded not to discuss the details of the cyber task, as per the nondisclosure agreement.

*12:00 P.M. to 4:00 P.M. (Cyber Task, Part 2):* Participants continued the network penetration task. Proctors continued to note the timing of breaks and any extreme behaviors in

HSR subjects.

*4:00 P.M. to 4:30 P.M. (Briefing):* All activity on the attack laptops was halted and participants filled out the Day 2 Red Team Briefing followed by the Overall Briefing (see Appendix A).

*4:30 P.M. to 5:15 P.M. (Task Battery or Report Writing):* Participants who opted out of the HSR portion continued to write a report on the cyber task (continuing the red team briefing). Participants who opted into the HSR portion complete the following tasks in order: Day 2 Cyber Task Questionnaire (hard copy), Deception Questionnaire (hard copy), Operation Span (computer), Need for Cognition (computer), Remote Associates Task (computer), Sandia Matrices (computer), Insight/Analytical Problem Solving (computer), and Sleep Quality Questionnaire (computer).

*5:15 P.M. to 5:30 P.M. (Wrap-Up):* Participants were debriefed and reminded not to discuss the task with others. Proctors handed out gift cards and collected the Empatica E4 devices from participants who participated in the HSR portion of the study.

## ACKNOWLEDGEMENTS

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525 SAND2018-5110 C.
Opinions expressed are those of the authors and not necessarily those of the U.S. Government.